

A Statistical Approach for Predicting Thermal Diffusivity Profiles in Fusion Plasmas as a Transport Model

Masayuki YOKOYAMA

National Institute for Fusion Science, Toki 509-5292, Japan

(Received 18 July 2014 / Accepted 11 August 2014)

A statistical approach is proposed to predict thermal diffusivity profiles as a transport “model” in fusion plasmas. It can provide regression expressions for the ion and electron heat diffusivities (χ_i and χ_e), separately, to construct their radial profiles. An approach that this letter is proposing outstrips the conventional scaling laws for the global confinement time (τ_E) since it also deals with profiles (temperature, density, heating depositions etc.). This approach has become possible with the analysis database accumulated by the extensive application of the integrated transport analysis suite to experiment data. In this letter, TASK3D-a [M. Yokoyama *et al.*, Plasma Fusion Res. **9**, 3402017 (2014)] analysis database for high-ion-temperature (high- T_i) plasmas [H. Takahashi *et al.*, Nucl. Fusion **53**, 073034 (2013)] in the LHD (Large Helical Device) [O. Kaneko *et al.*, Nucl. Fusion **53**, 104015 (2013)] is used as an example to describe an approach.

© 2014 The Japan Society of Plasma Science and Nuclear Fusion Research

Keywords: multiple ordinary least squares (OLS) regression analysis, TASK3D-a, LHD analysis database, energy transport

DOI: 10.1585/pfr.9.1302137

Conventionally, scaling laws for the global energy confinement time (τ_E) have been one of the approaches to systematically grasp the energy confinement of fusion plasmas [1, 2], and are also considered as one of the guidelines to design/predict future devices, such as ITER [3].

On the other hand, physics-based transport models have been employed to predict the plasma performance, such as expected temperature profiles for certain plasma operation scenarios. In such predictions, it has been always problematic whether employed transport model(s) are actually responsible for governing energy confinement in plasmas to be forecasted. In other words, how do we validate the employed transport models?

In this letter, a statistical approach is proposed to overcome such a problem. The goal of this approach is to predict thermal diffusivity profiles (in other words, transport “modelling”) based on the analysis database, without assuming any physics-based transport models. The accumulating and increasing analysis database should elucidate, by itself, the systematic dependence on plasma parameters. In some sense, the analysis database created by experimental power balance analysis can be considered to be the most “validated” transport model, since data are obtained from actual experimental data. This is the background of this letter.

It should be mentioned that a similar approach based on neural network has also been reported based on extensive case-analyses in DIII-D very recently [4].

Recent development of the integrated transport analysis suite, TASK3D-a (analysis version for LHD exper-

iments), and its extensive application to a wide-ranging LHD plasmas have created the analysis database which includes profile information such as ion and electron temperatures (T_i and T_e), electron density (n_e), NBI heating deposition, and ion and electron thermal diffusivities (χ_i and χ_e), and others.

TASK3D-a, in brief, consists of modules for temperature/density profile fittings, VMEC [5] equilibrium specification, NBI deposition calculations [6] and steady-state/dynamic energy transport calculations, so that they are sequentially executed in an automated manner [7]. Analyses for multiple-timings in multiple discharges form the analysis database. The employed data in this letter have been carefully checked in terms of completeness and measurement accuracy of kinetic profiles, reasonable profile fitting, and NBI injection without breakdown, and other criteria. The entire TASK3D-a database also includes data obtained by discharge scenario other than high- T_i discharges. However, high- T_i plasmas can provide wide range of T_i in a single shot, and this fact is essential to assure the statistical confidence. Thus, in this letter, selected small part of the analysis database (high- T_i plasmas) is utilized to describe an approach which this letter seeks to propose.

Figure 1 shows radial profiles of fitted values of (a) T_i (measured by charge exchange spectroscopy [8]), (b) T_e (Thomson scattering system [9]), and (c) the electron density, n_e (Thomson scattering system data calibrated with far-infrared laser interferometry [10]) of analysed cases. The number of discharges considered in this letter is 31. Multiple timings are analyzed in each discharge (corresponding to the timing of T_i -profile measurement, leading

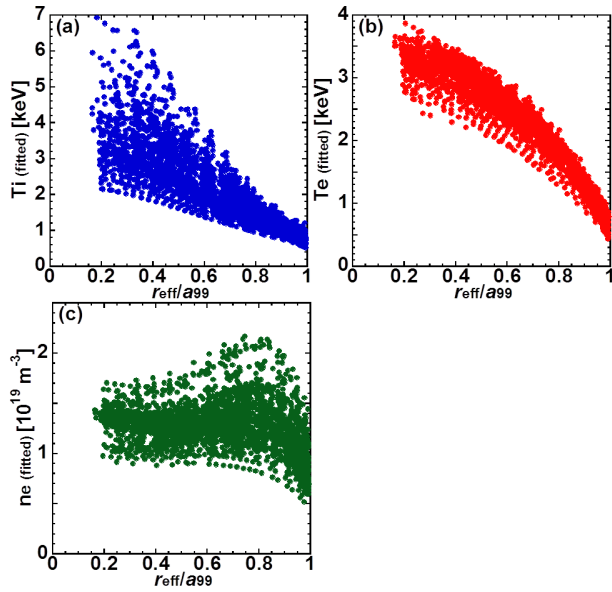


Fig. 1 Radial profiles (fitted) of (a) T_i , (b) T_e , and (c) n_e in analysis database.

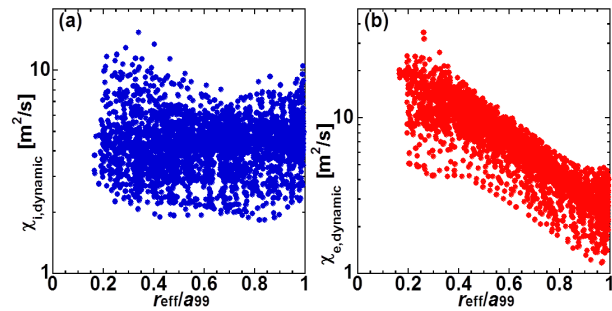


Fig. 2 Radial profiles of (a) χ_i and (b) χ_e in analysis database. The total number of data points of χ_i and χ_e is around 3,000, respectively.

to a total of about 200 timings), so that the evolution of T_i at the core region from low- T_i to high- T_i phase can be tracked, not only at the timing with the highest values of T_i . The T_e profile is rather stiff compared to that of T_i in this database. The n_e profiles are flat to hollow. Figure 2 shows (a) χ_i and (b) χ_e , obtained from the dynamic transport analysis [11] which takes into account the NBI slowing down and the temporal change of plasma parameters. The total number of data points (either ion or electron) shown in Fig. 2 is around 3,000.

The accumulation of TASK3D-a analyses results has led to the attempt at regression analysis for χ_i and χ_e with certain parameters as regression parameters. On performing regression analysis for the TASK3D-a analysis database, χ_i and χ_e are dimensionally normalized by Bohm diffusion coefficients, $T_{i,e}/(eB)$. Candidate predictor variables are also made into dimensionless variables, such as, for ions, the collision frequency normalized by that of the plateau_Pfirsch-Schlüter boundary (ν_i^*), the normalized

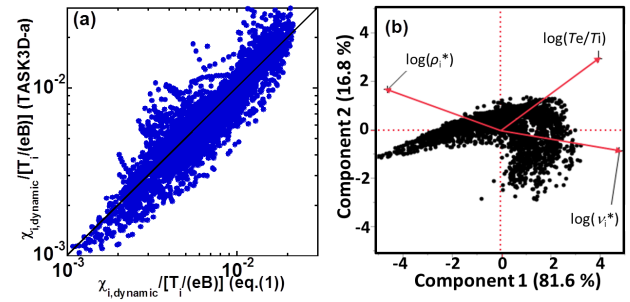


Fig. 3 (a) The comparison of $\chi_i/[T_i/(eB)]$ values between TASK3D-a analysis database and the regression results. (b) Biplot for predictor variables employed in Eq. (1).

Larmor radius (ρ_i^*), and the temperature ratio (T_e/T_i).

There are wide freedoms for the selection of the predictor variables including combinations of variables. In addition to the above mentioned three variables (ν_i^* , ρ_i^* , T_e/T_i), physically important variables such as E_r (radial electric field) shearing rate, and others do exist. However, the complete implementation of such variables into the analysis database has not yet been done. Thus, let me limit in this letter to proposing an approach that employs the available variables. Thus the results below are not necessarily the final and decisive conclusions.

Here, as a standard exercise in scaling studies, the assumed simple power-law scaling model has been transformed to the log-linear form. Multiple ordinary least squares (OLS) regression analysis has resulted in the regression expression for $\chi_i/[T_i/(eB)]$, by employing data shown in Fig. 2.

$$\chi_{i,\text{fit}}/[T_i/(eB)] = 6.08 \times 10^{-9} \nu_i^{*-0.139} \rho_i^{*-2.29} (T_e/T_i)^{0.77}. \quad (1)$$

Figure 3 (a) shows comparison of $\chi_i/[T_i/(eB)]$ values between TASK3D-a analysis database and the values predicted by Eq. (1). Around 3,000 data points, corresponding to a wide range of T_i and radial positions, are reasonably aligned on the diagonal line in Fig. 3 (a). The multicollinearities between predictor variables (dependencies in the multi-dimensional space) may have a destructive impact on the OLS regression analysis. However, it usually does not affect predictions provided the multi-collinearity remains similar. Nevertheless, an exact check for the multi-collinearity must be done before using the model. The best way to assess the multi-collinearity is applying principal component analysis (PCA) [12]. Figure 3 (b) shows a biplot [13, 14] (data (black dots), and eigenvectors (red arrows) of predictor variables) in the plane defined by the first two largest principal components (cf., Appendix). The angles less than 90 degrees between eigenvectors, which are seen in Fig. 3 (b), indicate there are dependencies between them, but still within acceptable range. An important statistical measure of the quality of the model is the ratio R^2 of the variation explained by the model to the

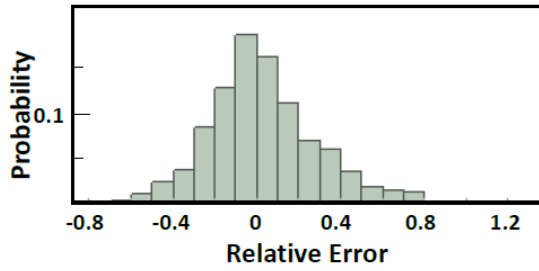


Fig. 4 The probability of the relative error, $(\chi_{i_eq.(1)} - \chi_{i_TASK3D-a})/\chi_{i_TASK3D-a}$, for the regression analysis on χ_i .

total variation. The obtained value is $R^2 = 0.83$, which is a relatively high value, indicating that Eq. (1) reasonably reproduces the response variable, that is, $\chi_i/[T_i/(eB)]$. The value of the root-mean-square-error (RMSE) is 0.28.

Figure 4 shows the probability distribution as a function of the relative error evaluated by $[(\chi_{i_eq.(1)} - \chi_{i_TASK3D-a})/\chi_{i_TASK3D-a}]$. Its mean value is about 3.7% with the standard deviation of about 28%. The relative error must be further reduced, of course. However, the conventional approach assuming physics-based models may not reach such a status for around 3,000 χ_i data. Furthermore, Eq. (1) covers the evolution of plasma parameters (from low- T_i to high- T_i as seen in Fig. 1 (a)) in high- T_i plasmas in the LHD. Thus, such a regression expression can be directly implemented into the predictive simulation instead of physics-based transport models. It should be noted here that the direct profile extrapolation approach proposed in Ref. [15] can also be the plausible method to predict profile evolution.

As for electrons, $\chi_e/[T_e/(eB)]$, the same set of predictor variables as ions, $(v_e^*, \rho_e^*, T_e/T_i)$, gives a regression expression with only $R^2 = 0.21$. This poor confidence level of statistics may be attributed to the small range of T_e (cf., Fig. 1 (b)) in the database causing a smaller range of predictor variables compared to that of ions. Increasing the number of predictor variables, from three to nine (addition of available variables such as the normalized temperature and density scale lengths, R/L_{Te} , R/L_{ne} , the rotational transform, the effective helicity [16], the dominant helicity and the toroidicity), can increase R^2 to 0.54. However, this value of R^2 is still rather small, and, moreover, configuration-related variables have high multicollinearity with each other as expected. Recently, trials have been conducted in the LHD experiment for increasing T_e in high- T_i plasmas (from $T_i > T_e$ towards $T_i \sim T_e$) by means of the increased available ECH power [17]. A corresponding increase of TASK3D-a analysis database (inclusion of higher T_e cases in high- T_i plasmas) is foreseen, when it is anticipated to increase the confidence level of statistics for electrons, as well.

Examples for comparisons of χ_i profiles between those deduced from Eq. (1) and from analysis database

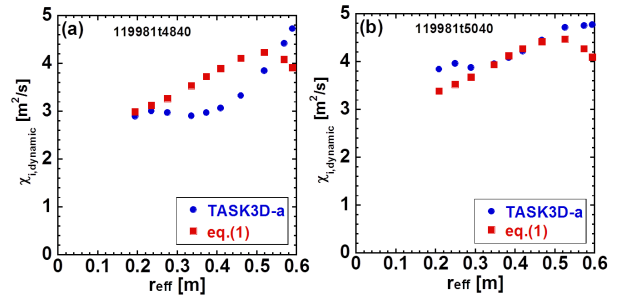


Fig. 5 Comparisons of χ_i profiles at two time slices, (a) 4.84 s and (b) 5.04 s of LHD shot number 119981. Circles (squares) symbols are TASK3D-a (obtained from Eq. (1)), respectively.

(TASK3D-a results) are shown in Fig. 5 at two time slices of LHD shot number 119981. It is found that in these two cases, χ_i data of analysis database (circle symbols) are reasonably well reproduced by deduced from Eq. (1). Of course, based on the existence of the relative error shown in Fig. 4, there are other cases without such a reasonable reproduction. Improving the statistical confidence such as with adding physically-important variables into regression analysis will be pursued.

In this letter, a statistical approach is proposed to predict thermal diffusivity profiles in fusion plasmas. The extensive application of the integrated transport analysis suite TASK3D-a to the LHD experiment has made this approach possible. A statistically-confident regression expression for the ion thermal diffusivity (covering wide range of T_i , also from core to edge) for LHD high- T_i plasmas has been provided, which can be directly implemented into the predictive simulation as a transport “model”.

It may be ultimately anticipated to elucidate the regression expression which is appropriate for ion and electron thermal diffusivities, separately, regardless of the confinement mode. Further application of TASK3D-a in a wider range of LHD plasmas and the resulting increase of analysis database will be performed in this direction.

Lastly, it should be emphasized that this approach is comprehensive for any other combinations of integrated transport analysis suites and fusion experiments.

Acknowledgments

The author is grateful to Dr. A. Kus (Max-Planck Institut für Plasmaphysik) for his extensive efforts for instructing statistical approach based on his long career on international fusion (tokamak and Stellarator-Heliotron) database. The author is also grateful to the LHD Experiment Group for making the experimental data available, and TASK3D-UD (users and developers) colleagues for making extensive transport analyses of NBI-heated LHD plasmas possible. In particularly, Dr. C. Suzuki is acknowledged for his cooperation on handling the analysis database so that it is suitable for statistical analysis. Advice

and encouragement from Prof. H. Sugama and Prof. K. Ida (National Institute for Fusion Science) have been highly appreciated. This work has been supported by the NIFS Collaborative Research Programs, NIFS11KNTT008 and NIFS11UNTT006, and by KEIN1107, NIFS/NINS (National Institutes of Natural Sciences) under the project, “Promotion of the International Collaborative Research Network Formation” for Coordinated Working Group Activity. The author also appreciates a grant-in-aid from the Future Energy Research Association (Kyoto).

Appendix. Supplementary Description for a Biplot (Fig. 3 (b)) [14]

The biplot is a graphical tool, based on the eigen-analysis, to investigate the structure of analyzed data. It is constructed as the projection of the dispersion onto a plane defined by the two selected principal components (PCs). The plot consists of both data points (shown as black dots) and eigenvectors (shown as red arrows) of predictor variables.

For collinearity checks, the eigen-analysis and PC analysis are usually made on the correlation matrix. In such a biplot, the cosine of the angle between arrows equals the correlation between the variables (angles close to 90 or 270 degrees indicate small correlation, angles of 0 or 180 degrees mean correlation of 1 or -1 , respectively). This information is most important for collinearity check.

For scaled and centered data, one can directly compare the lengths of arrows to see which variable are well represented in the selected plane (longer arrow means the variable lies near the plane). When necessary (the representation of a variable in the plane is not satisfactory), further biplots using other PC pairs should be created and

analysed. The distance between two points in the biplot corresponds the Euclidean distance between two observations in the multivariate space (dots far away from each other correspond to distant observations, and vice versa).

The Fig. 3 (b) shows a biplot in the plane PC1-PC2, accounting $(81.6 + 16.8 =) 98.4\%$ of the total variation (more than 80% of it belongs to the PC1). The angles between arrows are not too small, which indicates that dependencies between predictor variables are acceptable.

- [1] ITER Physics Expert Groups on Confinement and Transport and Confinement Modelling and Database *et al.*, Nucl. Fusion **39**, 2175 (1999).
- [2] H. Yamada *et al.*, Nucl. Fusion **45**, 1684 (2005).
- [3] <http://www.iter.org/>
- [4] O. Meneghini *et al.*, Phys. Plasmas **21**, 060702 (2014).
- [5] S.P. Hirshman and J.C. Whitson, Phys. Fluids **26**, 3553 (1983).
- [6] S. Murakami *et al.*, Trans. Fusion Technol. **27**, 256 (1995).
- [7] M. Yokoyama *et al.*, Plasma Fusion Res. **9**, 3402017 (2014).
- [8] M. Yoshinuma *et al.*, Fusion Sci. Technol. **58**, 375 (2010).
- [9] I. Yamada *et al.*, Fusion Sci. Technol. **58**, 345 (2010).
- [10] K. Tanaka *et al.*, Fusion Sci. Technol. **58**, 70 (2010).
- [11] H. Lee *et al.*, Plasma Phys. Control. Fusion **55**, 014011 (2013).
- [12] I.T. Jolliffe, *Principal Component Analysis, Springer Series in Statistics, 2nd ed.* (Springer, NY, 2002).
- [13] K.R. Gabriel, Biometrika **58**, 453 (1971).
- [14] J.O. Rawlings *et al.*, *Applied Regression Analysis: A Research Tool, Second Edition* (Springer, NY, 1998).
- [15] J. Miyazawa *et al.*, Fusion Eng. Des. **86**, 2879 (2011).
- [16] C.D. Beidler and W.N.G. Hitchon, Plasma Phys. Control. Fusion **36**, 317 (1994).
- [17] K. Nagaoka *et al.*, *to be presented* in the 25th IAEA Fusion Energy Conference, St. Petersburg, Oct. 2014.